

#6



RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/665,358

TIME: 09:58:24

Input Set : N:\Crf3\RULE60\09665358.raw

Output Set: N:\CRF4\09202002\I665358.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Stojiljkovic, Igor
 6 So, Magdalene
 7 Hwa, Vivian
 8 Heffron, Fred
 9 Nassif, Xavier

11 (ii) TITLE OF INVENTION: A Novel Bacterial Hemoglobin Receptor Gene and Uses

13 (iii) NUMBER OF SEQUENCES: 18

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Banner & Allegretti, Ltd.
 17 (B) STREET: 10 South Wacker Drive, Suite 3000
 18 (C) CITY: Chicago
 19 (D) STATE: Illinois
 20 (E) COUNTRY: USA
 21 (F) ZIP: 60606

ENTERED

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk
 25 (B) COMPUTER: IBM PC compatible
 26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/665,358
 C--> 31 (B) FILING DATE: 19-Sep-2000
 32 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: US/08/537,361
 36 (B) FILING DATE: 02 OCT 1995

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: Noonan, Kevin E
 41 (B) REGISTRATION NUMBER: 35,303
 42 (C) REFERENCE/DOCKET NUMBER: 94,784-A

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: 312-715-1000
 46 (B) TELEFAX: 312-715-1234
 47 (C) TELEX: 810-221-8317

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 2373 base pairs
 54 (B) TYPE: nucleic acid
 55 (C) STRANDEDNESS: single
 56 (D) TOPOLOGY: linear

58 (ii) MOLECULE TYPE: cDNA

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/665,358

TIME: 09:58:24

Input Set : N:\Crif3\RULE60\09665358.raw

Output Set: N:\CRF4\09202002\I665358.raw

```

60      (ix) FEATURE:
61          (A) NAME/KEY: CDS
62          (B) LOCATION: 1..2373
64      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
66 ATG AAA CCA TTA CAA ATG CTC CCT ATC GCC GCG CTG GTC GGC AGT ATT      48
67 Met Lys Pro Leu Gln Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile
68   1           5           10           15
70 TTC GGC AAT CCG GTC TTG GCA GCA GAT GAA GCT GCA ACT GAA ACC ACA      96
71 Phe Gly Asn Pro Val Leu Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr
72           20           25           30
74 CCC GTT AAG GCA GAG ATA AAA GCA GTG CGC GTT AAA GGT CAG CGC AAT      144
75 Pro Val Lys Ala Glu Ile Lys Ala Val Arg Val Lys Gly Gln Arg Asn
76           35           40           45
78 GCG CCT GCG GCT GTG GAA CGC GTC AAC CTT AAC CGT ATC AAA CAA GAA      192
79 Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu
80           50           55           60
82 ATG ATA CGC GAC AAT AAA GAC TTG GTG CGC TAT TCC ACC GAT GTC GGC      240
83 Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly
84   65           70           75           80
86 TTG AGC GAC AGC GGC CGC CAT CAA AAA GGC TTT GCT GTT CGC GGC GTG      288
87 Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val
88           85           90           95
90 GAA GGC AAC CGT GTC GGC GTG AGC ATA GAC GGT GTA AAC CTG CCT GAT      336
91 Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp
92           100          105          110
94 TCC GAA GAA AAC TCG CTG TAC GCC CGT TAT GGC AAC TTC AAC AGC TCG      384
95 Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser
96           115          120          125
98 CGT TTG TCT ATC GAC CCC GAA CTC GTA CGC AAT ATT GAA ATC GTG AAG      432
99 Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Glu Ile Val Lys
100          130          135          140
102 GGC GCA GAC TCT TTC AAT ACC GGC AGT GGT GCA TTG GGC GGC GGT GTG      480
103 Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val
104 145          150          155          160
106 AAT TAC CAA ACG CTG CAA GGC CGT GAT TTG CTG TTG GAC GAC AGG CAA      528
107 Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Asp Asp Arg Gln
108           165          170          175
110 TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC ACG CGT AAC CGT GAA TGG      576
111 Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp
112           180          185          190
114 ACA AAT ACC CTC GGT TTC GGT GTG AGT AAC GAC CGC GTG GAT GCT GCT      624
115 Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala
116           195          200          205
118 TTG CTG TAT TCG CAA CGG CGC GGC CAT GAA ACC GAA AGC GCG GGC AAC      672
119 Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Asn
120           210          215          220
122 CGC GGC TAT CCG GTA GAA GGT GCG GGT AAA GAA ACG AAT ATC CGC GGT      720
123 Arg Gly Tyr Pro Val Glu Gly Ala Gly Lys Glu Thr Asn Ile Arg Gly
124 225          230          235          240

```

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/665,358

TIME: 09:58:24

Input Set : N:\Crif3\RULE60\09665358.raw

Output Set: N:\CRF4\09202002\I665358.raw

126	TCC	GCC	CGC	GGC	ATC	CCC	GAT	CCG	TCC	AAA	CAC	AAA	TAC	CAC	AAC	TTC	768
127	Ser	Ala	Arg	Gly	Ile	Pro	Asp	Pro	Ser	Lys	His	Lys	Tyr	His	Asn	Phe	
128					245					250					255		
130	TTG	GGT	AAG	ATT	GCT	TAT	CAA	ATC	AAC	GAC	AAC	CAC	CGC	ATC	GGC	GCA	816
131	Leu	Gly	Lys	Ile	Ala	Tyr	Gln	Ile	Asn	Asp	Asn	His	Arg	Ile	Gly	Ala	
132					260					265					270		
134	TCG	CTC	AAC	GGT	CAG	CAG	GGG	CAT	AAT	TAC	ACG	GTT	GAA	GAG	TCT	TAT	864
135	Ser	Leu	Asn	Gly	Gln	Gln	Gly	His	Asn	Tyr	Thr	Val	Glu	Glu	Ser	Tyr	
136					275					280					285		
138	AAC	CTG	ACC	GCT	TCT	TCC	TGG	CGC	GAA	GCC	GAT	GAC	GTA	AAC	AGA	CGG	912
139	Asn	Leu	Thr	Ala	Ser	Ser	Trp	Arg	Glu	Ala	Asp	Asp	Val	Asn	Arg	Arg	
140					290					295					300		
142	CGC	AAT	GCC	AAC	CTC	TTT	TAC	GAA	TGG	ATG	CCT	GAT	TCA	AAT	TGG	TTG	960
143	Arg	Asn	Ala	Asn	Leu	Phe	Tyr	Glu	Trp	Met	Pro	Asp	Ser	Asn	Trp	Leu	
144	305					310						315				320	
146	TCG	TCT	TTG	AAG	GCG	GAC	TTC	GAT	TAT	CAG	AAA	ACC	AAA	GTG	GCG	GCG	1008
147	Ser	Ser	Leu	Lys	Ala	Asp	Phe	Asp	Tyr	Gln	Lys	Thr	Lys	Val	Ala	Ala	
148					325					330					335		
150	ATT	AAC	AAA	GGT	TCG	TTC	CCG	ACG	AAT	TAC	ACC	ACA	TGG	GAA	ACT	GAG	1056
151	Ile	Asn	Lys	Gly	Ser	Phe	Pro	Thr	Asn	Tyr	Thr	Thr	Trp	Glu	Thr	Glu	
152					340					345					350		
154	TAC	CAT	AAA	AAG	GAA	GTT	GGC	GAA	ATA	TAC	AAC	CGC	AGC	ATG	GAC	ACC	1104
155	Tyr	His	Lys	Lys	Glu	Val	Gly	Glu	Ile	Tyr	Asn	Arg	Ser	Met	Asp	Thr	
156					355					360					365		
158	CGA	TTC	AAA	CGT	TTT	ACT	TTG	CGT	TTG	GAC	AGC	CAT	CCG	TTG	CAA	CTC	1152
159	Arg	Phe	Lys	Arg	Phe	Thr	Leu	Arg	Leu	Asp	Ser	His	Pro	Leu	Gln	Leu	
160					370					375					380		
162	GGG	GGG	GGG	CGA	CAC	CGC	CTG	TCG	TTT	AAA	ACT	TTC	GCC	AGC	CGC	CGT	1200
163	Gly	Gly	Gly	Arg	His	Arg	Leu	Ser	Phe	Lys	Thr	Phe	Ala	Ser	Arg	Arg	
164	385					390						395				400	
166	GAT	TTT	GAA	AAC	CTA	AAC	CGC	GAC	GAT	TAT	TAC	TTC	AGC	GGC	CGT	GTT	1248
167	Asp	Phe	Glu	Asn	Leu	Asn	Arg	Asp	Asp	Tyr	Tyr	Phe	Ser	Gly	Arg	Val	
168					405					410					415		
170	GTT	CGA	ACC	ACC	AGC	AGT	ATC	CAG	CAT	CCG	GTG	AAA	ACC	ACC	AAC	TAC	1296
171	Val	Arg	Thr	Thr	Ser	Ser	Ile	Gln	His	Pro	Val	Lys	Thr	Thr	Asn	Tyr	
172					420					425					430		
174	GGT	TTC	TCA	CTG	TCT	GAC	CAA	ATT	CAA	TGG	AAC	GAC	GTG	TTC	AGT	AGC	1344
175	Gly	Phe	Ser	Leu	Ser	Asp	Gln	Ile	Gln	Trp	Asn	Asp	Val	Phe	Ser	Ser	
176					435					440					445		
178	CGC	GCA	GGT	ATC	CGT	TAC	GAC	CAC	ACC	AAA	ATG	ACG	CCT	CAG	GAA	TTG	1392
179	Arg	Ala	Gly	Ile	Arg	Tyr	Asp	His	Thr	Lys	Met	Thr	Pro	Gln	Glu	Leu	
180					450					455					460		
182	AAT	GCC	GAG	TGT	CAT	GCT	TGT	GAC	AAA	ACA	CCA	CCT	GCA	GCC	AAC	ACT	1440
183	Asn	Ala	Glu	Cys	His	Ala	Cys	Asp	Lys	Thr	Pro	Pro	Ala	Ala	Asn	Thr	
184	465					470						475				480	
186	TAT	AAA	GGC	TGG	AGC	GGT	TTT	GTC	GGC	TTG	GCG	GCG	CAA	CTG	AAT	CAG	1488
187	Tyr	Lys	Gly	Trp	Ser	Gly	Phe	Val	Gly	Leu	Ala	Ala	Gln	Leu	Asn	Gln	
188					485					490					495		
190	GCT	TGG	CGT	GTC	GGT	TAC	GAC	ATT	ACT	TCC	GGC	TAC	CGT	GTC	CCC	AAT	1536

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/665,358

TIME: 09:58:24

Input Set : N:\Crif3\RULE60\09665358.raw

Output Set: N:\CRF4\09202002\I665358.raw

191	Ala	Trp	Arg	Val	Gly	Tyr	Asp	Ile	Thr	Ser	Gly	Tyr	Arg	Val	Pro	Asn	
192				500					505					510			
194	GCG	TCC	GAA	GTG	TAT	TTC	ACT	TAC	AAC	CAC	GGT	TCG	GGT	AAT	TGG	CTG	1584
195	Ala	Ser	Glu	Val	Tyr	Phe	Thr	Tyr	Asn	His	Gly	Ser	Gly	Asn	Trp	Leu	
196			515					520					525				
198	CCC	AAT	CCC	AAC	CTG	AAA	GCC	GAG	CGC	AGC	ACC	ACC	CAC	ACC	CTG	TCT	1632
199	Pro	Asn	Pro	Asn	Leu	Lys	Ala	Glu	Arg	Ser	Thr	Thr	His	Thr	Leu	Ser	
200		530					535					540					
202	CTG	CAA	GGC	CGC	AGC	GAA	AAA	GGC	ATG	CTG	GAT	GCC	AAC	CTG	TAT	CAA	1680
203	Leu	Gln	Gly	Arg	Ser	Glu	Lys	Gly	Met	Leu	Asp	Ala	Asn	Leu	Tyr	Gln	
204	545					550				555						560	
206	AGC	AAT	TAC	CGC	AAT	TTC	CTG	TCT	GAA	GAG	CAG	AAG	CTG	ACC	ACC	AGC	1728
207	Ser	Asn	Tyr	Arg	Asn	Phe	Leu	Ser	Glu	Glu	Gln	Lys	Leu	Thr	Thr	Ser	
208				565					570					575			
210	GGC	ACT	CCC	GGC	TGT	ACT	GAG	GAA	AAT	GCT	TAC	TAC	AGT	ATA	TGC	AGC	1776
211	Gly	Thr	Pro	Gly	Cys	Thr	Glu	Glu	Asn	Ala	Tyr	Tyr	Ser	Ile	Cys	Ser	
212			580					585					590				
214	GAC	CCC	TAC	AAA	GAA	AAA	CTG	GAT	TGG	CAG	ATG	AAA	AAT	ATC	GAC	AAG	1824
215	Asp	Pro	Tyr	Lys	Glu	Lys	Leu	Asp	Trp	Gln	Met	Lys	Asn	Ile	Asp	Lys	
216			595				600					605					
218	GCC	AGA	ATC	CGC	GGT	ATC	GAG	CTG	ACA	GGC	CGT	CTG	AAT	GTG	GAC	AAA	1872
219	Ala	Arg	Ile	Arg	Gly	Ile	Glu	Leu	Thr	Gly	Arg	Leu	Asn	Val	Asp	Lys	
220		610					615					620					
222	GTA	GCG	TCT	TTT	GTT	CCT	GAG	GGC	TGG	AAA	CTG	TTC	GGC	TCG	CTG	GGT	1920
223	Val	Ala	Ser	Phe	Val	Pro	Glu	Gly	Trp	Lys	Leu	Phe	Gly	Ser	Leu	Gly	
224	625					630					635					640	
226	TAT	GCG	AAA	AGC	AAA	CTG	TCG	GGC	GAC	AAC	AGC	CTG	CTG	TCC	ACA	CAG	1968
227	Tyr	Ala	Lys	Ser	Lys	Leu	Ser	Gly	Asp	Asn	Ser	Leu	Leu	Ser	Thr	Gln	
228				645					650					655			
230	CCG	CTG	AAA	GTG	ATT	GCC	GGT	ATC	GAC	TAT	GAA	AGT	CCG	AGC	GAA	AAA	2016
231	Pro	Leu	Lys	Val	Ile	Ala	Gly	Ile	Asp	Tyr	Glu	Ser	Pro	Ser	Glu	Lys	
232				660					665					670			
234	TGG	GGC	GTA	TTC	TCC	CGC	CTG	ACC	TAT	CTG	GGC	GCG	AAA	AAG	GTC	AAA	2064
235	Trp	Gly	Val	Phe	Ser	Arg	Leu	Thr	Tyr	Leu	Gly	Ala	Lys	Lys	Val	Lys	
236			675					680					685				
238	GAC	GCG	CAA	TAC	ACC	GTT	TAT	GAA	AAC	AAG	GGC	TGG	GGT	ACG	CCT	TTG	2112
239	Asp	Ala	Gln	Tyr	Thr	Val	Tyr	Glu	Asn	Lys	Gly	Trp	Gly	Thr	Pro	Leu	
240		690					695					700					
242	CAG	AAA	AAG	GTA	AAA	GAT	TAC	CCG	TGG	CTG	AAC	AAG	TCG	GCT	TAT	GTG	2160
243	Gln	Lys	Lys	Val	Lys	Asp	Tyr	Pro	Trp	Leu	Asn	Lys	Ser	Ala	Tyr	Val	
244	705					710					715					720	
246	TTC	GAT	ATG	TAC	GGC	TTC	TAC	AAA	CCG	GTG	AAA	AAC	CTG	ACC	CTG	CGT	2208
247	Phe	Asp	Met	Tyr	Gly	Phe	Tyr	Lys	Pro	Val	Lys	Asn	Leu	Thr	Leu	Arg	
248				725					730					735			
250	GCG	GGC	GTG	TAC	AAC	CTG	TTC	AAC	CGC	AAA	TAC	ACC	ACT	TGG	GAT	TCC	2256
251	Ala	Gly	Val	Tyr	Asn	Leu	Phe	Asn	Arg	Lys	Tyr	Thr	Thr	Trp	Asp	Ser	
252				740					745					750			
254	CTG	CGC	GGT	TTA	TAT	AGC	TAC	AGC	ACC	ACC	AAT	GCG	GTC	GAC	CGC	GAT	2304
255	Leu	Arg	Gly	Leu	Tyr	Ser	Tyr	Ser	Thr	Thr	Asn	Ala	Val	Asp	Arg	Asp	

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/665,358

TIME: 09:58:24

Input Set : N:\Crif3\RULE60\09665358.raw

Output Set: N:\CRF4\09202002\I665358.raw

```

256          755          760          765
258 GGC AAA GGC TTA GAT CGC TAC CGC GCC CCA GGC CGC AAT TAC GCC GTA      2352
259 Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Gly Arg Asn Tyr Ala Val
260          770          775          780
262 TCG CTG GAA TGG AAG TTT TAA      2373
263 Ser Leu Glu Trp Lys Phe *
264 785          790
268 (2) INFORMATION FOR SEQ ID NO: 2:
270 (i) SEQUENCE CHARACTERISTICS:
271 (A) LENGTH: 790 amino acids
272 (B) TYPE: amino acid
273 (D) TOPOLOGY: linear
275 (ii) MOLECULE TYPE: protein
277 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
279 Met Lys Pro Leu Gln Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile
280 1 5 10 15
282 Phe Gly Asn Pro Val Leu Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr
283 20 25 30
285 Pro Val Lys Ala Glu Ile Lys Ala Val Arg Val Lys Gly Gln Arg Asn
286 35 40 45
288 Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu
289 50 55 60
291 Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly
292 65 70 75 80
294 Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val
295 85 90 95
297 Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp
298 100 105 110
300 Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser
301 115 120 125
303 Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Glu Ile Val Lys
304 130 135 140
306 Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val
307 145 150 155 160
309 Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Asp Asp Arg Gln
310 165 170 175
312 Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp
313 180 185 190
315 Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala
316 195 200 205
318 Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Asn
319 210 215 220
321 Arg Gly Tyr Pro Val Glu Gly Ala Gly Lys Glu Thr Asn Ile Arg Gly
322 225 230 235 240
324 Ser Ala Arg Gly Ile Pro Asp Pro Ser Lys His Lys Tyr His Asn Phe
325 245 250 255
327 Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala
328 260 265 270
330 Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr

```

VERIFICATION SUMMARY

DATE: 09/20/2002

PATENT APPLICATION: US/09/665,358

TIME: 09:58:25

Input Set : N:\Crf3\RULE60\09665358.raw

Output Set: N:\CRF4\09202002\I665358.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]